

Specification

Antitumor Protein and Gene Encoding Same

5 BACKGROUND OF THE INVENTION

Field of the Invention

The present invention relates to an antitumor protein and an nucleotide sequence encoding the same.

Background Art

10 Various studies have been conducted on antitumor substances found in edible mushrooms.

For example, polysaccharides and glycoprotein from mushrooms which have antitumor activity are disclosed in Japanese Patent Laid-open Publication Nos.61214/1977, 15 74797/1980, 293923/1986, 70362/1993 and 80699/1994, Japanese Patent Publication Nos. 47518/1986, 47519/1986 and 26172/1991. It is also reported that mushrooms are found to have antitumor activity when administered.

However, there has been no report on an amino acid 20 sequence of an antitumor protein derived from Tricholoma matsutake which directly kills a tumor cell and on a gene encoding said protein.

SUMMARY OF THE INVENTION

The inventors now have purified an antitumor protein 25 derived from Tricholoma matsutake and determined an amino acid sequence as well as a cDNA sequence encoding the protein. Further, the inventors have successfully purified the cDNA sequence and obtained a recombinant antitumor protein expressed in E. coli which is transformed by 30 introducing a vector comprising the cDNA sequence. The present invention is based on these findings.

Thus, an object of the present invention is to provide an antitumor protein, a fragment of said protein, a nucleotide molecule encoding said protein, a vector 35 comprising said molecule, a host cell transformed by said vector, a process for preparing said protein, and an antibody against said protein.

The protein according to the present invention comprises

- (a) an amino acid sequence of SEQ ID No.1, or
- (b) a modified amino acid sequence of SEQ ID No.1 which has antitumor activity wherein one or more amino acids are added and/or inserted into the amino acid sequence of SEQ ID No.1 and/or one or more amino acids in the amino acid sequence of SEQ ID No.1 are substituted and/or deleted.

The protein according to the present invention is useful as an antitumor agent.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates the structure of plasmid vector pTS18.

Figure 2 illustrates the deletion of the TTM gene. Dashed lines denote a deleted area.

DETAILED DESCRIPTION OF THE INVENTION

Protein

The protein according to the present invention comprises the amino acid sequence of SEQ ID No.1. A protein consisting of the amino acid sequence of SEQ ID No.1 has antitumor activity as described in examples.

Examples of the proteins according to the present invention include those consisting of a modified amino acid sequence of SEQ ID No.1 which has antitumor activity wherein one or more amino acids are added and/or inserted into the amino acid sequence of SEQ ID NO.1 and/or one or more amino acids in the amino acid sequence of SEQ ID NO.1 are substituted and/or deleted. The terms "addition", "insertion", "substitution" and "deletion" refer to those which do not damage the antitumor activity of the protein consisting of the amino acid sequence of SEQ ID NO.1. The numbers of modifications such as additions, insertions, substitutions and deletions may be in the range between 1 and 8.

An addition, insertion, substitution or deletion may be introduced into an amino acid sequence in accordance with, for example, Molecular Cloning (A laboratory manual),

second edition, Cold Spring Harbor Laboratory Press, Vol. 2, Chap. 15 (1989); Botstein, D. et al., Science, 229:1193 (1985); Craik, C.S., Bio. Techniques, 3:12 (1985); Itakura, K. et al., Annu. Rev. Biochem. 53:323 (1984); Shortle, D. et al., Annu. Rev. Genet. 15:265 (1981); or Smith, M. Annu. Rev. Genet. 19:423 (1985).

The wording "protein which have antitumor activity" as used herein refers to a protein which is evaluated by one skilled in the art to have antitumor activity, for example, a protein which is evaluated to have antitumor activity as tested under the conditions in Example 1 (3).

The molecular weight of the protein consisting of the amino acid sequence of SEQ ID NO.1 is about 65 kDa as measured by SDS-PAGE.

The amino acid sequence of SEQ ID NO.1 can be prepared by expression of the DNA sequence of SEQ ID NO.2 in a bacteria using a common technique. The cDNA sequence can be prepared by screening a cDNA library derived from Tricholoma matsutake using an antibody against the antitumor protein as a probe (see Example 2).

The protein according to the present invention has antitumor activity. Therefore, the protein according to the present invention may be formulated in a pharmaceutical composition which is used in the treatment of tumor such as carcinoma of uterine cervix or corpus uteri, and a variety of cancers caused by abnormal expression of antioncogene p53 or pBR (e.g., carcinoma cutaneum, lung cancer, liver cancer, kidney cancer, and breast cancer).

The pharmaceutical composition according to the present invention may be administered to a mammal including a human perorally or parenterally (e.g., intramuscularly, intravenously, subcutaneously, intrarectally, percutaneously or pernasally) in a form suitable for peroral or parenteral administration. A formulation which directly reaches a target area (e.g., a tablet which dissolves at a specific site, a liniment, or an injection) may be preferably used in the treatment of tumor.

The protein according to the present invention may be formulated in oral drugs (e.g., tablets, capsules, granules, powder, pills, grains, troches) when considering stability of the protein and the drug delivery path; injectable drugs (e.g., for intravenous or intramuscular injection); intrarectal drugs; and soluble or insoluble suppositories depending on its intended use. The pharmaceutical composition in these forms may be prepared by conventional methods with pharmaceutically acceptable vehicles such as bulking agents and fillers; adjuvants such as binding agents, wetting agents, disintegrants, surfactants, lubricants, dispersers, buffering agents, and solution adjuvant; additives such as preservatives, antiseptics, flavouring agents, soothing agents, stabilizers, colouring agents, and sweetening agents. A dose for various treatments may be determined depending on the route of administration as well as the age, sex, and condition of the patient.

Nucleotide Sequence

The present invention provides a nucleotide sequence encoding the protein according to the present invention. Examples of such nucleotide sequences include those comprising all or part of the DNA sequence of SEQ ID NO.2. Other examples of such nucleotide sequences include those comprising all or part of the DNA sequence in SEQ ID NO.2.

As mentioned above, the DNA sequence of SEQ ID NO.2 was obtained from a cDNA library derived from Tricholoma matsutake. This DNA sequence contains an open reading frame of the protein which starts at ATG (1-3) and ends at TAA (1699-1701).

The amino acid sequence determines a number of possible base sequences that encode the amino acid sequence in SEQ ID NO.1.

When the amino acid sequence of the protein according to the present invention is given, a nucleotide sequence encoding the amino acid sequence is easily determined, and a variety of nucleotide sequences encoding the amino acid

sequence of SEQ ID NO: 1 can be selected.

Thus, a nucleotide sequence encoding the protein according to the present invention include DNA sequences which degenerate as a result of the genetic code as to the DNA sequence of SEQ ID No.2 as well as RNA sequences corresponding to the DNA sequences.

The nucleotide sequence according to the present invention may be naturally occurred or obtained by synthesis. It may also be synthesized with a part of a sequence derived from the naturally occurring one. DNAs may typically be obtained by screening a chromosome library or a cDNA library in accordance with conventional methods in the field of genetic engineering, for example, by screening a chromosome library or a cDNA library with an appropriate DNA probe obtained based on information of the partial amino acid sequence. The nucleotide sequence according to the present invention can be prepared, for example, from Tricholoma matsutake cDNA library by using an oligonucleotide encoding a peptide selected from SEQ ID Nos.3-18 as a screening probe.

The nucleotide sequences from nature are not specifically restricted to any sources; but may be derived from Tricholoma matsutake or other sources.

Vectors and Transformed Cells

The present invention provides a vector comprising the nucleotide sequence according to the present invention in such a manner that the vector can be replicable and express the protein encoded by the nucleotide sequence in a host cell. In addition, according to the present invention, we provide a host cell transformed by the vector. There is no other restriction to the host-vector system. It may express proteins fused with other proteins. Examples of an expression system of a fusion protein include those expressing MBP (maltose binding protein), GST (glutathione-S-transferase), HA (hemagglutinin), polyhistidine, myc, and Fas.

Examples of such systems expressing fusion proteins

include those expressing β -galactosidase, glutathione-S-transferase, and luciferase.

Examples of vectors include plasmid vectors (e.g., pBluescript SK(-), pBluescript SK(+), pGEX-4T, pGEX-5T, 5 pRIT2T, pBPV, and pSVK3 (Pharmacia, etc.); ZAP Express, pYEUra3, pMAM, and pOG (Toyobo); pET-11a, b, c, and d, pET-20b, pET-28a, b, and c, and pET-32a and b (Novagen); pQE-10, 16, 30, 40, 50, 60, and 70 (Qiagen); virus vectors (e.g., retrovirus vectors and adenovirus vectors); and 10 liposome vectors (e.g., cationic liposome vectors).

In order to prepare a desired protein in the host cell, the vector according to the present invention may have a sequence which regulate expression of the protein (e.g., a promoter sequence, a terminator sequence, or an enhancer 15 sequence) or markers for selecting a host cell (e.g., a neomycin-resistant gene or a kanamycin-resistant gene). Further, the vector may have the nucleotide sequence according to the present invention in a repeated form (e.g., in a tandem form). Such additional sequences may be 20 introduced into the vector. A host cell may be transformed by the vector by conventional methods.

The vector according to the present invention may be prepared by conventional methods and procedures of the genetic engineering field.

25 Examples of host cells include E. coli (e.g., SOLR, JM109, XL1-Blue MRF', and BL21(DE3)), yeast cells (e.g., YRG-2), Bacillus subtilis, animal cells (e.g., CHO cells, COS cells, human keratinocytes, COP-5, C127, mouse 3T3 cells, FR3T3, and HB101).

30 The protein according to the present invention is obtained from the culture by culturing host cells which are transformed as described above in an appropriate medium. Therefore, the present invention provides a process for preparing the protein according to the present invention. 35 Such a process enables mass production of an antitumor protein.

The culture of the transformed host cell and culture

condition may essentially be the same as those for the cell to be used. In addition, the protein according to the present invention may be recovered from the culture medium and purified according to conventional methods, for example, chromatography such as ion exchange chromatography, gel filtration chromatography, and immunoaffinity chromatography

Antibody

The present invention provides an antibody against the protein according to the present invention. The term "antibody" as used herein includes a polyclonal antibody or a monoclonal antibody.

The antibody according to the present invention can be prepared by conventional methods, for example, by injecting the protein of SEQ ID NO.1 or a fragment thereof into an animal (e.g., rabbit, rat or mouse) together with suitable carriers (e.g., Freund's complete and incomplete adjuvants) and then purifying the serum from the animal after a certain period.

Specific reaction (i.e., immuno reaction) of the antibody may be used as an indicator of an antitumor protein. Therefore, the antibody according to the present invention may be used for purifying and screening an antitumor protein.

Examples

The present invention is further illustrated by the following Examples which are not intended as a limitation of the invention.

Example 1 Purification of Antitumor Protein

(1) Purification of Protein

An antitumor protein was purified from commercially available (or wild) fresh Tricholoma matsutake by homogenizing it in accordance with conventional methods and then isolating using purifying procedures such as column chromatography, HPLC, and electrophoresis. The detailed procedure is as follows:

A Tris buffer solution containing NaCl and protease

inhibitor (50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM PMSF, 1 mM EDTA, 0.1 mM IAA (iodoacetamide), 1 µg/ml pepstatin A, and 1 µg/ml leupeptin) was used for the preliminary elution of the protein, followed by precipitation with ammonium sulfate (90% saturated ammonium sulfate). The precipitate was dialyzed with 25 mM Tris-HCl (pH 7.5) containing 1/10 the above protease inhibitor (PI) to desalt. Then, after DEAE Toyopearl (ion exchange chromatography), concentration of the active fraction, purification through phenyl Sepharose (hydrophobic chromatography), concentration of the active fraction, gel filtration by HPLC (TSK gel G3000SW), the purified protein was finally obtained.

In ion exchange chromatography and hydrophobic chromatography, 25 mM Tris-HCl (pH 7.5) containing PI was used as eluant. For linear concentration gradient, NaCl and $(\text{NH}_4)_2\text{SO}_4$ were used, respectively. In gel filtration, 0.1 M sodium phosphate buffer (pH 7.2) containing 0.1 M Na_2SO_4 and PI was used as eluant.

The sample obtained by gel filtration with HPLC was analyzed by SDS-PAGE. The protein on the gel, which was transferred to a PVDF membrane and stained with CBB, exhibited a single band (about 65 kDa).

It was found that when Tricholoma matsutake with no freshness was used or when no protease inhibitor was used in purifying procedures, yield and antitumor activity were found to be lower.

Some of the samples were recovered by staining the gel with CBB after SDS-PAGE, and cutting it to extract electrically. These samples were used to determine the amino acid sequence (Example 2).

It was also found that the protein can be purified by affinity chromatography using a column in which the antibody (see (2)) was bound to CNBr-activated Sepharose 6MB resin (Pharmacia).

(2) Polyclonal Antibody

A rabbit was immunized with the protein purified in (1) to prepare antiserum. The procedure is as follows:

The purified protein, 15 µg, was mixed with Freund's complete adjuvant, stirred intensely to emulsion, and subcutaneously injected to the back of a rabbit. After 3 weeks, the rabbit was boosted with 150 µg of the purified protein, which was mixed with Freund's incomplete adjuvant to give emulsion. Then, after 2 weeks, they were directly reboosted using 50 µg of antibody, and blood was collected from its earlobe 1 week later.

Next, 5 ml of antiserum was incubated at 56°C for 30 min, mixed with 5 ml of PBS(-) and the same amount of saturated $(\text{NH}_4)_2\text{SO}_4$, and maintained still in iced water. After centrifugation, the precipitate was redissolved in sodium phosphate buffer solution and mixed with an additional amount of saturated $(\text{NH}_4)_2\text{SO}_4$ to a final $(\text{NH}_4)_2\text{SO}_4$ concentration of 20%. After centrifugation, the supernatant was recovered and mixed with an additional amount of saturated $(\text{NH}_4)_2\text{SO}_4$ to a final $(\text{NH}_4)_2\text{SO}_4$ concentration of 33%. After centrifugation, the precipitate was recovered and redissolved. It was then dialyzed and desalted, followed by ion exchange chromatography (DE52 resin), to give an IgG fraction.

(3) Antitumor Activity Test

Lethal activity was investigated on cells which had been transformed by simian virus 40 (SV40) and human papiloma virus (HPV) which were known to cause malignant alteration. More specifically, antitumor activity was estimated with lethal activity. When the protein purified in (1) above was given to the cells, the quantity of the tested protein necessary for 50% fatal activity of total cells was 10 ng/ml in SVT2 cells (transformed SV40 cells), 100 ng/ml in A31 cells (transformed SV40 cells), and 15-20 ng/ml in human preputial cells (transformed HPV16 cells).

Example 2 cDNA Cloning and Sequencing

The amino acid sequence at the N-terminal of the

protein purified in Example 1 was determined (SEQ ID NOS.3 and 4) using a protein sequencer (Hewlett-Packard).

Also, the protein obtained in Example 1 was digested using lysyl endonuclease to give a number of peptide fragments. Among them, the amino acid sequences of 14 peptide fragments were determined (SEQ ID NOS.5-18).

On the other hand, Tricholoma matsutake mRNA was purified with oligo-dT Latex (oligo-dT particles; Takara), then with STRATAGENE ZAP-cDNA Synthesis Kit (available from Toyobo), to synthesize cDNA. After synthesized, the cDNA was packaged in vitro in lambda phage using Gigapak III Gold (Stratagene, available from Toyobo) to prepare a phage library.

Using the antibody obtained in Example 1 (2) as a probe, the phage library was screened for the antitumor gene. Twenty-one phages were tested positive. The procedure is as follows:

The concentration of the library was determined with titer. About 2,000 to 20,000 phages and 600µl E. coli (XL1-Blue) were plated in 150 mm NZYM culture plates together with 6ml NZYM Top Agar (0.7%). They were incubated at 42°C for 3-4 hours until plaques developed to suitable sizes of about 1 mm. Then, a 130-140 mm nitrocellulose membrane soaked with 10 mM IPTG was placed on each plate, and incubation was continued at 37°C for 3 hours. After the plates were cooled at 4°C for 1 hour or more, the nitrocellulose filters were removed from the plates, and shaken in TBS-T buffer solution containing 3% skim milk.

Next, the filters were soaked in the buffer solution of the primary antibody (Example 1 (2)), and gently shaken in TBS-T buffer solution containing 3% skim milk. The filters were then soaked in the buffer solution of secondary antibody conjugated to alkali phosphatase (AP), and washed with TBS-T buffer solution. After they were washed with alkali phosphatase (AP) buffer solution, positive phages were detected.

The resulting positive phages were transformed with SOLR strains (Stratagene) by in vivo excision, using ZAP-cDNA Synthesis Kit (available from Toyobo) according to a manufacturer's manual.

5 Plasmid pTS18 as shown in Fig. 1 was obtained from the transformants. Plasmid pTS18 (containing the cDNA sequence in SEQ ID NO. 1) was used in Example 3 as an expression vector.

10 The resulting pTS18 was deleted by using Exo/Mung DNA Sequencing System (Stratagene), blunted at both terminals, and ligated with self-DNA (Fig. 2). Next, E. coli JM109 (Toyobo) was transformed with the deleted plasmid DNA. The nucleotide sequences of the portions of the gene into which deletion mutation was introduced were completely determined
15 using ABI PRISM Cycle Sequencing Kit (Parkin Elmer) both on the sense and anti-sense chains.

The determined partial sequences were used to establish the complete amino acid sequence and cDNA sequence (SEQ ID NO.2) of the antitumor protein. A deduced molecular weight
20 was about 62 kDa. The amino acid sequence on the N terminal (SEQ ID NOS.3 and 4) agreed with the amino acid sequence 2-30 and the amino acid sequence 2-58 in SEQ ID NO.1.

Also, the sequences of the peptide fragments (SEQ ID
25 NOS.5-18) agreed with the amino acid sequence in SEQ ID NO.1 as follows:

SEQ ID NO.5: 59-77 in SEQ ID NO.1;
SEQ ID NO.6: 89-149 in SEQ ID NO.1;
SEQ ID NO.7: 150-178 in SEQ ID NO.1;
30 SEQ ID NO.8: 179-209 in SEQ ID NO.1;
SEQ ID NO.9: 210-267 in SEQ ID NO.1;
SEQ ID NO.10: 268-297 in SEQ ID NO.1;
SEQ ID NO.11: 298-355 in SEQ ID NO.1;
SEQ ID NO.12: 356-406 in SEQ ID NO.1;
35 SEQ ID NO.13: 407-436 in SEQ ID NO.1;
SEQ ID NO.14: 437-486 in SEQ ID NO.1;
SEQ ID NO.15: 487-521 in SEQ ID NO.1;

SEQ ID NO.16: 522-554 in SEQ ID NO.1;

SEQ ID NO.17: 555-566 in SEQ ID NO.1;

SEQ ID NO.18: 78-99 in SEQ ID NO.1.

5 These peptide fragments are useful as antigens for
obtaining an antibody against the antitumor protein which
can be used in a method for screening and purifying an
antitumor protein.

Example 3 Production of Antitumor Protein (1)

10 Competent cells (JM109 strain; Toyobo) stored at -80°C
were melted, and 100 µl of the cells was transferred to
Falcon tube (code 2059). It was mixed with deleted clones
of pTS18 (Example 2) and allowed to stand in iced water for
30 min. After exposed to a thermal shock (42°C) for 30 s,
15 it was cooled in ice for 2 min. After 900 µl SOC culture
was added, it was incubated at 37°C for 1 hour with
shaking. The cells were then planted in an LB/Amp plate
in an appropriate amount, and incubated overnight at 37°C.
A colony having an area of a platinum ring that appeared
on the plate was transplanted to a liquid LB culture
20 (containing Amp), and incubated at 37°C until absorption
at 660 nm (Abs660) increased to about 0.2. Then, after
IPTG was added to a final concentration of 10 mM, the
culture was incubated until Abs660 increased to about 1.

25 The cells were suspended in the extract (50 mM
Tris-HCl, pH 7.5) used in Example 1 (1), which contained
PI, and ultrasonically destroyed. After the extract (50
mM Tris-HCl) was centrifuged, the supernatant was recovered
in the eluate via affinity chromatography (CNBr-activated
30 Sepharose 6MB resin; Pharmacia) binding the antibody
described in Example 1 (2).

The eluate was analyzed by SDS-PAGE combined with
Western blotting using the antibody described in Example
1 (2). The result showed that the protein according to the
present invention was expressed in the host cell.

35 Example 4 Production of Antitumor Protein (2)

(1) Preparation of expression vector pET-28a

A DNA fragment encoding the antitumor protein was generated by polymerase chain reaction (PCR) using plasmid pTS18 (10 ng)(Example 2) as a template DNA. PCR reaction was carried out using reagents packaged in a commercially available kit (TAKARA Co.) and the following primers (5 pmole, each) in accordance with a manufacturer's manual.

Primer 1:GAGAGACCATGGGGTATCGTCTTTCC (SEQ ID NO.19)

Primer 2:GAGAGAGGATCCGGAGACGCCAAGGAT (SEQ ID NO.20)

After the PCR reaction, the product was digested by NcoI and BamHI. The resulting fragment (0.1 μ g) was ligated into the NcoI/BamHI site of pET-28a (0.5 μ g) (Novagen).

The resulting DNA construct was introduced into competent cells (E. Coli, DH5 α and JM109 strains; Toyobo). The plasmid DNA which was harvested from the transformed cells was introduced into competent cells (BL21 (DE3) strain; Novagen).

(2) Preparation of expression vector pET-28b

A DNA fragment encoding the antitumor protein was prepared by digesting plasmid pTS18 (Example 2) by EcoRI and XhoI and collecting EcoRI/XhoI fragments. The resulting fragments (0.1 μ g) were ligated into the EcoRI/XhoI site of pET-28b (0.5 μ g)(Novagen).

The resulting DNA construct was introduced into competent cells (E. Coli, DH5 α and JM109 strains; Toyobo). The plasmid DNA which was harvested from the transformed cells was introduced into competent cells (BL21 (DE3) strain; Novagen).

(3) Expression of antitumor protein gene

One loopful of the transformed cells, BL21 (DE3) strain having pET-28a and BL21 (DE3) strain having pET-28b, obtained as described in Example 3 (1) and (2) were inoculated on 1 ml of NZYM medium containing 50 μ g/ml of Kanamycin and preincubated at 37°C overnight. 100 μ l taken from the cultured medium was inoculated on 10 ml of NZYM medium containing 50 μ g/ml of kanamycin and incubated at 25°C until Abs600 increased to about 0.4. After IPTG was

added to a final concentration of 1.0 mM, the culture was incubated for 24 hours.

5 The cells were harvested from the culture medium, suspended in the extract (25 mM Tris-HCl, pH 7.0) used in Example 1 (1) containing PI, and ultrasonically destroyed.

After the extract (25 mM Tris-HCl, pH 7.0) was centrifuged, the precipitate was recovered. The precipitate was analyzed by SDS-PAGE. A single band was observed on the position of 65 kDa.

10 The precipitate was also analyzed by Western blotting using the antibody described in Example 1 (2). An immunoreactive band was observed at the same position as that observed on the SDS-PAGE gel. This result showed that the gene of the antitumor protein was expressed in the host
15 cells.

SEQUENCE LENGTH: 556

STRANDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20	Ser	Gln	Gly	Asp	Leu	Thr	Ser	Ser	Gln	His	Glu	Ile	Val	His	Phe	Thr
				20				25						30		
	Asp	Val	Phe	Ile	Ala	Gly	Ser	Gly	Pro	Ile	Ser	Cys	Thr	Tyr	Ala	Arg
				35				40						45		
	His	Ile	Ile	Asp	Asn	Thr	Ser	Thr	Thr	Lys	Val	Tyr	Met	Ala	Glu	Ile
				50				55						60		
	Gly	Ser	Gln	Asp	Asn	Pro	Val	Ile	Gly	Ala	His	His	Lys	Asn	Ser	Ile
				65				70						80		
	Lys	Phe	Gln	Lys	Asp	Ile	Asp	Lys	Phe	Val	Asn	Ile	Ile	Asn	Gly	Ala
				85				90						95		
	Leu	Gln	Pro	Ile	Ser	Ile	Ser	Pro	Ser	Asp	Thr	Tyr	Gln	Pro	Thr	Leu
				100				105						110		
	Ala	Val	Ala	Ala	Trp	Ala	Pro	Pro	Ile	Asp	Pro	Ala	Glu	Gly	Gln	Leu
				115				120						125		
	Val	Ile	Met	Gly	His	Asn	Pro	Asn	Gln	Glu	Ala	Gly	Leu	Asn	Leu	Pro
				130				135						140		
	Gly	Ser	Ala	Val	Thr	Arg	Thr	Val	Gly	Gly	Met	Ala	Thr	His	Trp	Thr
				145				150						160		
	Cys	Ala	Cys	Pro	Thr	Pro	His	Asp	Glu	Glu	Arg	Val	Asn	Asn	Pro	Val
				165				170						175		
Asp	Lys	Gln	Glu	Phe	Asp	Ala	Leu	Leu	Glu	Arg	Ala	Lys	Thr	Leu	Leu	
			180				185						190			
Asn	Val	His	Ser	Asp	Gln	Tyr	Asp	Asp	Ser	Ile	Arg	Gln	Ile	Val	Val	
			195				200						205			
Lys	Glu	Thr	Leu	Gln	Gln	Thr	Leu	Asp	Ala	Ser	Arg	Gly	Val	Thr	Thr	
			210				215						220			

Leu	Pro	Leu	Gly	Val	Glu	Arg	Arg	Thr	Asp	Asn	Pro	Ile	Tyr	Val	Thr	225	230	235	240
Trp	Thr	Gly	Ala	Asp	Thr	Val	Leu	Gly	Asp	Val	Pro	Lys	Ser	Pro	Arg	245	250	255	
Phe	Ala	Leu	Val	Thr	Glu	Thr	Arg	Val	Thr	Lys	Leu	Ile	Val	Ser	Glu	260	265	270	
Thr	Asn	Pro	Thr	Gln	Val	Val	Ala	Ala	Leu	Leu	Arg	Asn	Leu	Asn	Thr	275	280	285	
Ser	Asn	Asp	Glu	Leu	Val	Val	Ala	Lys	Ser	Phe	Val	Ile	Ala	Cys	Gly	290	295	300	
Ala	Val	Cys	Thr	Pro	Gln	Ile	Leu	Trp	Asn	Ser	Asn	Ile	Arg	Pro	Tyr	305	310	315	320
Ala	Leu	Gly	Arg	Tyr	Leu	Ser	Glu	Gln	Ser	Met	Thr	Phe	Cys	Gln	Ile	325	330	335	
Val	Leu	Lys	Arg	Gly	Ile	Val	Asp	Ala	Ile	Ala	Thr	Asp	Pro	Arg	Phe	340	345	350	
Ala	Ala	Lys	Val	Glu	Ala	His	Lys	Lys	Lys	His	Pro	Asp	Asp	Val	Leu	355	360	365	
Pro	Ile	Pro	Phe	His	Glu	Pro	Glu	Pro	Gln	Val	Met	Ile	Pro	Tyr	Thr	370	375	380	
Ser	Asp	Phe	Pro	Trp	His	Val	Gln	Val	His	Arg	Asp	Ala	Phe	Ser	Tyr	385	390	395	400
Gly	Asp	Val	Gly	Pro	Lys	Ala	Asp	Pro	Arg	Val	Val	Val	Asp	Leu	Arg	405	410	415	
Phe	Phe	Gly	Lys	Ser	Asp	Ile	Val	Glu	Glu	Asn	Arg	Val	Thr	Phe	Gly	420	425	430	
Pro	Asn	Pro	Lys	Leu	Arg	Glu	Trp	Glu	Ala	Gly	Val	Thr	Asp	Thr	Tyr	435	440	445	
Gly	Met	Pro	Gln	Pro	Thr	Phe	His	Val	Lys	Arg	Thr	Asn	Ala	Asp	Gly	450	455	460	
Asp	Arg	Asp	Gln	Arg	Met	Met	Asn	Asp	Met	Thr	Asn	Val	Ala	Asn	Met	465	470	475	480
Leu	Gly	Gly	Tyr	Leu	Pro	Gly	Ser	Tyr	Pro	Gln	Phe	Met	Ala	Pro	Gly	485	490	495	
Leu	Val	Leu	His	Ile	Thr	Gly	Thr	Thr	Arg	Ile	Gly	Thr	Asp	Asp	Gln	500	505	510	


```

Thr Ser Val Ala Asp Pro Thr Ser Lys Val His Asn Phe Asn Asn Leu
      515                      520                      525
Trp Val Gly Gly Asn Gly Cys Ile Pro Asp Ala Thr Ala Cys Asn Pro
      530                      535                      540
Thr Arg Thr Ser Val Ala Tyr Ala Leu Lys Gly Ala Glu Ala Val Val
545                      550                      555                      560
Asn Tyr Leu Gly Val Ser      *
                        565

```

SEQ ID NO.2

SEQUENCE LENGTH: 1701

SEQUENCE TYPE: nucleic acid

STRANDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to RNA

SEQUENCE DESCRIPTION

```

ATG CCG ATA CGT CTT TCC AAA GAA AAA ATC AAC GAC CTG CTG CAA CGT      48
Met Pro Ile Arg Leu Ser Lys Glu Lys Ile Asn Asp Leu Leu Gln Arg
1           5           10           15
TCT CAA GGG GAT CTT ACT TCC TCG CAA CAC GAA ATT GTA CAT TTC ACT      96
Ser Gln Gly Asp Leu Thr Ser Ser Gln His Glu Ile Val His Phe Thr
           20           25           30
GAT GTT TTC ATT GCT GGC AGT GGT CCC ATT AGC TGT ACT TAC GCC CGC      144
Asp Val Phe Ile Ala Gly Ser Gly Pro Ile Ser Cys Thr Tyr Ala Arg
           35           40           45
CAC ATC ATT GAC AAT ACC TCA ACT ACA AAG GTT TAC ATG GCC GAA ATA      192
His Ile Ile Asp Asn Thr Ser Thr Thr Lys Val Tyr Met Ala Glu Ile
           50           55           60
GGT TCT CAA GAT AAC CCT GTC ATC GGG GCC CAT CAC AAG AAC TCC ATA      240
Gly Ser Gln Asp Asn Pro Val Ile Gly Ala His His Lys Asn Ser Ile
           65           70           75           80
AAG TTT CAG AAA GAC ATT GAC AAG TTT GTG AAT ATC ATC AAC GGT GCC      288
Lys Phe Gln Lys Asp Ile Asp Lys Phe Val Asn Ile Ile Asn Gly Ala
           85           90           95

```

CTC	CAG	CCG	ATT	TCG	ATT	TCG	CCA	TCG	GAC	ACC	TAC	CAG	CCC	ACT	CTC	336
Leu	Gln	Pro	Ile	Ser	Ile	Ser	Pro	Ser	Asp	Thr	Tyr	Gln	Pro	Thr	Leu	
			100					105					110			
GCT	GTA	GCA	GCG	TGG	GCG	CCG	CCC	ATC	GAT	CCT	GCC	GAA	GGC	CAG	CTC	384
Ala	Val	Ala	Ala	Trp	Ala	Pro	Pro	Ile	Asp	Pro	Ala	Glu	Gly	Gln	Leu	
		115					120					125				
GTG	ATT	ATG	GGA	CAC	AAT	CCG	AAT	CAG	GAG	GCC	GGC	CTG	AAC	CTT	CCC	432
Val	Ile	Met	Gly	His	Asn	Pro	Asn	Gln	Glu	Ala	Gly	Leu	Asn	Leu	Pro	
	130					135					140					
GGT	AGC	GCT	GTC	ACT	AGG	ACA	GTC	GGG	GGG	ATG	GCG	ACC	CAC	TGG	ACT	480
Gly	Ser	Ala	Val	Thr	Arg	Thr	Val	Gly	Gly	Met	Ala	Thr	His	Trp	Thr	
145					150					155					160	
TGC	GCG	TGT	CCT	ACT	CCA	CAT	GAC	GAA	GAG	AGG	GTC	AAC	AAC	CCA	GTT	528
Cys	Ala	Cys	Pro	Thr	Pro	His	Asp	Glu	Glu	Arg	Val	Asn	Asn	Pro	Val	
			165						170				175			
GAC	AAG	CAG	GAG	TTC	GAC	GCA	CTG	CTC	GAA	CGT	GCT	AAA	ACA	TTG	CTC	576
Asp	Lys	Gln	Glu	Phe	Asp	Ala	Leu	Leu	Glu	Arg	Ala	Lys	Thr	Leu	Leu	
		180					185					190				
AAC	GTT	CAC	AGC	GAC	CAG	TAC	GAC	GAT	TCT	ATC	CGT	CAG	ATA	GTT	GTC	624
Asn	Val	His	Ser	Asp	Gln	Tyr	Asp	Asp	Ser	Ile	Arg	Gln	Ile	Val	Val	
	195						200					205				
AAA	GAG	ACT	CTT	CAG	CAG	ACC	CTT	GAT	GCG	TCG	CGG	GGT	GTG	ACC	ACT	672
Lys	Glu	Thr	Leu	Gln	Gln	Thr	Leu	Asp	Ala	Ser	Arg	Gly	Val	Thr	Thr	
	210					215					220					
CTC	CCG	CTG	GGG	GTG	GAG	CGC	CGT	ACG	GAC	AAT	CCT	ATT	TAT	GTC	ACC	720
Leu	Pro	Leu	Gly	Val	Glu	Arg	Arg	Thr	Asp	Asn	Pro	Ile	Tyr	Val	Thr	
225					230				235					240		
TGG	ACC	GGT	GCC	GAT	ACC	GTC	CTT	GGT	GAT	GTG	CCG	AAG	AGT	CCC	CGA	768
Trp	Thr	Gly	Ala	Asp	Thr	Val	Leu	Gly	Asp	Val	Pro	Lys	Ser	Pro	Arg	
		245						250				255				
TTC	GCT	TTG	GTT	ACA	GAG	ACG	AGA	GTG	ACG	AAG	CTT	ATT	GTC	AGT	GAA	816
Phe	Ala	Leu	Val	Thr	Glu	Thr	Arg	Val	Thr	Lys	Leu	Ile	Val	Ser	Glu	
		260					265				270					
ACC	AAT	CCG	ACG	CAG	GTT	GTT	GCT	GCG	TTG	CTA	CGT	AAC	TTG	AAT	ACA	864
Thr	Asn	Pro	Thr	Gln	Val	Val	Ala	Ala	Leu	Leu	Arg	Asn	Leu	Asn	Thr	
	275						280				285					

AGC	AAC	GAT	GAA	CTT	GTC	GTG	GCC	AAG	AGT	TTC	GTC	ATA	GCT	TGT	GGA	912
Ser	Asn	Asp	Glu	Leu	Val	Val	Ala	Lys	Ser	Phe	Val	Ile	Ala	Cys	Gly	
290							295				300					
GCA	GTC	TGC	ACA	CCG	CAA	ATC	TTG	TGG	AAC	AGC	AAC	ATC	CGC	CCA	TAT	960
Ala	Val	Cys	Thr	Pro	Gln	Ile	Leu	Trp	Asn	Ser	Asn	Ile	Arg	Pro	Tyr	
305					310				315						320	
GCG	CTT	GGT	CGC	TAC	CTC	AGC	GAA	CAG	TCC	ATG	ACT	TTT	TGT	CAG	ATC	1008
Ala	Leu	Gly	Arg	Tyr	Leu	Ser	Glu	Gln	Ser	Met	Thr	Phe	Cys	Gln	Ile	
			325						330					335		
GTT	CTC	AAG	AGG	GGC	ATA	GTC	GAT	GCC	ATC	GCT	ACT	GAC	CCT	CGC	TTC	1056
Val	Leu	Lys	Arg	Gly	Ile	Val	Asp	Ala	Ile	Ala	Thr	Asp	Pro	Arg	Phe	
		340						345				350				
GCT	GCG	AAG	GTT	GAG	GCG	CAC	AAG	AAG	AAG	CAC	CCC	GAT	GAC	GTG	CTG	1104
Ala	Ala	Lys	Val	Glu	Ala	His	Lys	Lys	Lys	His	Pro	Asp	Asp	Val	Leu	
		355					360				365					
CCC	ATT	CCA	TTC	CAC	GAG	CCT	GAA	CCT	CAA	GTG	ATG	ATT	CCG	TAC	ACG	1152
Pro	Ile	Pro	Phe	His	Glu	Pro	Glu	Pro	Gln	Val	Met	Ile	Pro	Tyr	Thr	
		370					375				380					
TCG	GAC	TTC	CCT	TGG	CAT	GTT	CAG	GTG	CAT	CGC	GAT	GCA	TTC	TCA	TAT	1200
Ser	Asp	Phe	Pro	Trp	His	Val	Gln	Val	His	Arg	Asp	Ala	Phe	Ser	Tyr	
385					390					395					400	
GGT	GAT	GTT	GGA	CCC	AAG	GCC	GAC	CCG	CGT	GTT	GTC	GTC	GAT	CTG	AGG	1248
Gly	Asp	Val	Gly	Pro	Lys	Ala	Asp	Pro	Arg	Val	Val	Val	Asp	Leu	Arg	
			405						410					415		
TTT	TTC	GGC	AAA	TCA	GAT	ATT	GTC	GAA	GAA	AAT	CGA	GTG	ACT	TTC	GGT	1296
Phe	Phe	Gly	Lys	Ser	Asp	Ile	Val	Glu	Glu	Asn	Arg	Val	Thr	Phe	Gly	
		420					425				430					
CCG	AAC	CCT	AAG	CTA	CGC	GAG	TGG	GAA	GCG	GGT	GTT	ACA	GAC	ACT	TAT	1344
Pro	Asn	Pro	Lys	Leu	Arg	Glu	Trp	Glu	Ala	Gly	Val	Thr	Asp	Thr	Tyr	
		435					440				445					
GGA	ATG	CCA	CAG	CCG	ACA	TTC	CAT	GTC	AAG	CGG	ACC	AAC	GCC	GAT	GGA	1392
Gly	Met	Pro	Gln	Pro	Thr	Phe	His	Val	Lys	Arg	Thr	Asn	Ala	Asp	Gly	
		450				455					460					
GAC	CGT	GAC	CAG	AGG	ATG	ATG	AAT	GAT	ATG	ACC	AAC	GTC	GCG	AAC	ATG	1440
Asp	Arg	Asp	Gln	Arg	Met	Met	Asn	Asp	Met	Thr	Asn	Val	Ala	Asn	Met	
465					470					475					480	

CTG	GGT	GGG	TAC	CTT	CCT	GGC	TCC	TAC	CCT	CAA	TTT	ATG	GCA	CCT	GGT	1488
Leu	Gly	Gly	Tyr	Leu	Pro	Gly	Ser	Tyr	Pro	Gln	Phe	Met	Ala	Pro	Gly	
				485					490					495		
CTC	GTA	CTG	CAC	ATC	ACG	GGA	ACT	ACT	CGG	ATC	GGG	ACA	GAT	GAT	CAA	1536
Leu	Val	Leu	His	Ile	Thr	Gly	Thr	Thr	Arg	Ile	Gly	Thr	Asp	Asp	Gln	
			500					505					510			
ACT	TCT	GTT	GCT	GAT	CCG	ACA	TCA	AAG	GTT	CAT	AAC	TTC	AAC	AAT	CTG	1584
Thr	Ser	Val	Ala	Asp	Pro	Thr	Ser	Lys	Val	His	Asn	Phe	Asn	Asn	Leu	
			515				520					525				
TGG	GTC	GGC	GGG	AAT	GGG	TGC	ATT	CCA	GAT	GCG	ACT	GCC	TGC	AAC	CCG	1632
Trp	Val	Gly	Gly	Asn	Gly	Cys	Ile	Pro	Asp	Ala	Thr	Ala	Cys	Asn	Pro	
			530			535					540					
ACT	CGT	ACG	AGC	GTC	GCG	TAT	GCG	CTC	AAG	GGT	GCT	GAG	GCT	GTA	GTC	1680
Thr	Arg	Thr	Ser	Val	Ala	Tyr	Ala	Leu	Lys	Gly	Ala	Glu	Ala	Val	Val	
			545			550			555					560		
AAT	TAC	CTT	GGC	GTC	TCC	TGA										1701
Asn	Tyr	Leu	Gly	Val	Ser	*										
				565												

SEQ ID NO.3

SEQUENCE LENGTH: 29

SEQUENCE TYPE: amino acid

STRANDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION

Pro Ile Arg Leu Ser Lys Glu Lys Ile Asn Asp Leu Leu Gln Arg Ser

1

5

10

15

Gln Gly Asp Leu Thr Ser Ser Gln His Glu Ile Val His

20

25

SEQ ID NO.4

SEQUENCE LENGTH: 57

SEQUENCE TYPE: amino acid

STRANDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION

Pro Ile Arg Leu Ser Lys Glu Lys Ile Asn Asp Leu Leu Gln Arg Ser
 1 5 10 15
 Gln Gly Asp Leu Thr Ser Ser Gln His Glu Ile Val His Phe Thr Asp
 20 25 30
 Val Phe Ile Ala Gly Ser Gly Pro Ile Ser Cys Thr Tyr Ala Arg His
 35 40 45
 Ile Ile Asp Asn Thr Ser Thr Thr Lys
 50 55

SEQ ID NO.5

SEQUENCE LENGTH: 19

SEQUENCE TYPE: amino acid

STRANDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION

Val Tyr Met Ala Glu Ile Gly Ser Gln Asp Asn Pro Val Ile Gly Ala
 1 5 10 15
 His His Lys

SEQ ID NO.6

SEQUENCE LENGTH: 61

SEQUENCE TYPE: amino acid

STRANDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION

Phe	Val	Asn	Ile	Ile	Asn	Gly	Ala	Leu	Gln	Pro	Ile	Ser	Ile	Ser	Pro
1				5					10					15	
Ser	Asp	Thr	Tyr	Gln	Pro	Thr	Leu	Ala	Val	Ala	Ala	Trp	Ala	Pro	Pro
			20				25						30		
Ile	Asp	Pro	Ala	Glu	Gly	Gln	Leu	Val	Ile	Met	Gly	His	Asn	Pro	Asn
			35				40						45		
Gln	Glu	Ala	Gly	Leu	Asn	Leu	Pro	Gly	Ser	Ala	Val	Thr			
	50					55					60				

SEQ ID NO.7

SEQUENCE LENGTH: 29

SEQUENCE TYPE: amino acid

STRANDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION

Arg	Thr	Val	Gly	Gly	Met	Ala	Thr	His	Trp	Thr	Cys	Ala	Cys	Pro	Thr
1					5				10					15	
Pro	His	Asp	Glu	Glu	Arg	Val	Asn	Asn	Pro	Val	Asp	Lys			
			20					25							

SEQ ID NO.8

SEQUENCE LENGTH: 31

SEQUENCE TYPE: amino acid

STRANDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION

Gln	Glu	Phe	Asp	Ala	Leu	Leu	Glu	Arg	Ala	Lys	Thr	Leu	Leu	Asn	Val
1					5					10					15
His	Ser	Asp	Gln	Tyr	Asp	Asp	Ser	Ile	Arg	Gln	Ile	Val	Val	Lys	
				20					25					30	

SEQ ID NO.9

SEQUENCE LENGTH: 58

SEQUENCE TYPE: amino acid

STRANDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION

Glu	Thr	Leu	Gln	Gln	Thr	Leu	Asp	Ala	Ser	Arg	Gly	Val	Thr	Thr	Leu
1					5					10					15
Pro	Leu	Gly	Val	Glu	Arg	Arg	Thr	Asp	Asn	Pro	Ile	Tyr	Val	Thr	Trp
				20					25					30	
Thr	Gly	Ala	Asp	Thr	Val	Leu	Gly	Asp	Val	Pro	Lys	Ser	Pro	Arg	Phe
			35					40					45		
Ala	Leu	Val	Thr	Glu	Thr	Arg	Val	Thr	Lys						
			50					55							

SEQ ID NO.10

SEQUENCE LENGTH: 30

SEQUENCE TYPE: amino acid

STRANDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION

Leu	Ile	Val	Ser	Glu	Thr	Asn	Pro	Thr	Gln	Val	Val	Ala	Ala	Leu	Leu
1				5					10					15	
Arg	Asn	Leu	Asn	Thr	Ser	Asn	Asp	Glu	Leu	Val	Val	Ala	Lys		
			20					25					30		

SEQ ID NO.11

SEQUENCE LENGTH: 58

SEQUENCE TYPE: amino acid

STRANDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION

Ser	Phe	Val	Ile	Ala	Cys	Gly	Ala	Val	Cys	Thr	Pro	Gln	Ile	Leu	Trp
1				5					10					15	
Asn	Ser	Asn	Ile	Arg	Pro	Tyr	Ala	Leu	Gly	Arg	Tyr	Leu	Ser	Glu	Gln
			20					25				30			
Ser	Met	Thr	Phe	Cys	Gln	Ile	Val	Leu	Lys	Arg	Gly	Ile	Val	Asp	Ala
			35					40				45			
Ile	Ala	Thr	Asp	Pro	Arg	Phe	Ala	Ala	Lys						
			50				55								

SEQ ID NO.12

SEQUENCE LENGTH: 51

SEQUENCE TYPE: amino acid

STRANDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION

Val	Glu	Ala	His	Lys	Lys	Lys	His	Pro	Asp	Asp	Val	Leu	Pro	Ile	Pro
1				5				10						15	
Phe	His	Glu	Pro	Glu	Pro	Gln	Val	Met	Ile	Pro	Tyr	Thr	Ser	Asp	Phe
			20					25						30	
Pro	Trp	His	Val	Gln	Val	His	Arg	Asp	Ala	Phe	Ser	Tyr	Gly	Asp	Val
			35					40						45	
Gly	Pro	Lys													
			50												

SEQ ID NO.13

SEQUENCE LENGTH: 30

SEQUENCE TYPE: amino acid

STRANDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION

Ala	Asp	Pro	Arg	Val	Val	Val	Asp	Leu	Arg	Phe	Phe	Gly	Lys	Ser	Asp
1				5				10						15	
Ile	Val	Glu	Glu	Asn	Arg	Val	Thr	Phe	Gly	Pro	Asn	Pro	Lys		
			20					25					30		

SEQUENCE LENGTH: 50

STRANDNESS: single

MOLECULE TYPE: peptide

Leu Arg Glu Trp Glu Ala Gly Val Thr Asp Thr Tyr Gly Met Pro Gln

1 5 10 15

Pro Thr Phe His Val Lys Arg Thr Asn Ala Asp Gly Asp Arg Asp Gln

20 25 30

Arg Met Met Asn Asp Met Thr Asn Val Ala Asn Met Leu Gly Gly Tyr

30 40 45

Leu Pro

50

SEQUENCE LENGTH: 35

STRANDNESS: single

MOLECULE TYPE: peptide

Gly Ser Tyr Pro Gln Phe Met Ala Pro Gly Leu Val Leu His Ile Thr

1 5 10 15

Gly Thr Thr Arg Ile Gly Thr Asp Asp Gln Thr Ser Val Ala Asp Pro

20 25 30

Thr Ser Lys

35

SEQUENCE LENGTH: 33

STRANDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION

Val His Asn Phe Asn Asn Leu Trp Val Gly Gly Asn Gly Cys Ile Pro

1

5

10

15

Asp Ala Thr Ala Cys Asn Pro Thr Arg Thr Ser Val Ala Tyr Ala Leu

20

25

30

Lys

SEQUENCE LENGTH: 12

STRANDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION

Gly Ala Glu Ala Val Val Asn Tyr Leu Gly Val Ser

1

5

10

SEQUENCE LENGTH: 22

STRANDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION

Asn Ser Ile Lys Phe Gln Lys Asp Ile Asp Lys Phe Val Asn Ile Ile

1

5

10

15

Asn Gly Ala Leu Gln Pro

20

SEQ ID NO.19

SEQUENCE LENGTH: 26

SEQUENCE TYPE: nucleic acid

STRANDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: synthetic DNA

SEQUENCE DESCRIPTION

GAGAGACCAT GGGGTATCGT CTTTCC

26

SEQ ID NO.20

SEQUENCE LENGTH: 27

SEQUENCE TYPE: nucleic acid

STRANDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: synthetic DNA

SEQUENCE DESCRIPTION

GAGAGAGGAT CCGGAGACGC CAAGGAT

27

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125